Sampling principles for biodiversity study

Xubin Pan

1 **Sampling principles for biodiversity study**

2 Xubin Pan

3 Institute of Plant Quarantine, Chinese Academy of Inspection and Quarantine, Beijing 100029,

4 China

5 E-mail: xubin.hu.pan@gmail.com

6

7 **Abstract**

8 Sampling is a fundamental tool in ecology and critical for biodiversity measurement. However,

9 basic principles of biodiversity sampling have been overlooked for many years. In this paper, I

10 proposed and explored five principles of sampling for a specific area and biodiversity study.

11 The first principle of sampling, *species increasing with area*, is that the number of species

12 increases with the area. The second principle of sampling, *individuals increasing with area*, is

13 that the number of individuals increases with the area. The third principle of sampling, *sum of*

14 *species number*, is that the sum of species number in one area and species number in another

15 area is no less than the total species number in the two areas. The fourth principle of sampling,

16 *individual complement*, is that the sum of the mathematical expectation of individual number

17 of one or several species in the area *a* and that of the same one or several species in the area

18 *A-a* is the total individual number *N* of the same one or several species in the total area *A*. The

19 fifth principle of sampling, *species-area theory*, is that the sum of the mathematical expectation

20 of number of species in the area *a* and that of number of species lost if area *A-a* is cleared is the

21 total species number *M* in the total area *A*.

22 Keywords: Individuals-area relationship, Species-area relationship, Endemics-area relationship

23 **Introduction**

24 Sampling is a fundamental tool in ecology. Because the complete investigation of one

25 ecological question maybe involves a large area and long time, it will cost too much time and

26 money, and leads to infeasible in practice. Thus, sampling is unavoidable in the ecology study.

27 However, due to the complexity of ecosystem, such as interactions among organisms and

28 environment and their large spatial-temporal heterogeneity, how to sampling should be very

29 carefully, which will decide how to analyze the data from the sampling and the credibility of

30 the results. Thus, solid sampling theory is very important for ecology, although it didn't get too

31 much concern recently. As the core of the ecology, the spatial relationship between the species

32 and the area can be a good start for rethink the sampling theory.

33

34 Individual–area relationship (IAR), species-area relationship (SAR) and endemics-area

35 relationship (EAR) are important concepts in biodiversity conservation and habitat preservation

36 (WCMC, 1992; Kinzig and Harte, 2000; Connor et al., 2000; Green and Ostling, 2003; Chris et

37 al., 2004; World Resources Institute, 2005). However, there is still much debating over the

38 estimation of extinction rate based on SAR, which tends to make overestimation when

39 compared with observed extinction (Pimm and Askins, 1995; Rosenzweig, 1995; Harte and

40 Kinzig, 1997; He and Hubbell, 2011; Pan, 2013). One explanation is that such estimation

41 includes certain species that are "committed to extinction" instead of going extinct right away

42 after habitat clearing (Heywood, 1994; Tilman, 1994; Mace et al., 2003). Recently, He and

43 Hubbell (2011) suggested that the reason was the difference of between sampling areas based

44 SAR and EAR, because a sample area of a species for extinction is often larger than a sample

45   area of the same species for existence. Among these controversies, the core problem is the

46   sampling of biodiversity measurement, the basic principles of which have been overlooked for

47   many years.

48

49   However, much work has been focused on statistical or mathematical calculation based on

50   SAR, rather than the biological and ecological implication, especially the basic principles for

51   biodiversity sampling (Turner and Tjørve, 2005). Are there common laws for sampling in

52   biodiversity measurement? Here, we proposed five sampling principles for biodiversity study

53   in a specific area and the last two were proved, focusing on the change of species number and

54   individuals for one or several species with a changing area. This analysis will be helpful for the

55   establishment of theory platform for biodiversity and other ecological discussion.

56

57   **Theoretical Frame**

58   Biodiversity not just study the number of species, but also the amount of individuals. For the

59   sampling problems of biodiversity, is the relationship between the area and the number of

60   species and amount of individuals. Usually, it is thought the relationship between the number

61   of species and the area is saturation curve, and the relationship between the amount of

62   individuals and the area is increasing curve. The shape of this curve is influenced by the spatial

63   distribution and sampling collection/statistic.

64

65   **Sampling principles for biodiversity measurement**

66   The first principle of sampling, species increasing with area, is that the number of species

67   increases with the area. Assume the number of species is m in the area a, when the sampling

68   area increases from a to a', the number of species in area a' is m', and $m' \geq m$. If no new

69   species emerges in the area $a' - a$, $m' = m$.

70

71   The second principle of sampling, individuals increasing with area, is that the number of

72   individuals increases with the area. Assume the number of individuals is n in the area a,

73   when the sampling area increases from a to a', the number of individuals in area a' is m',

74   and $n' \geq n$. If there no new individuals emerges in the area $a' - a$, $n' = n$.

75

76   The third principle of sampling, sum of species number, is that the sum of species number in

77   one area and species number in another area is no less than the total species number in the two

78   areas. Assume the number of species is m_1 in the area a_1, the number of species is m_2 in the

79   area a_2, the number of species in area $a_1 + a_2$ is m', and $m' \leq m_1 + m_2$. If there is no

80   overlapping species between areas a_1 and a_2, $m' = m_1 + m_2$.

81

82   The fourth principle of sampling, individual complement, is that the sum of the mathematical

83   expectation of individual number of one or several species in area a and that of individual

84   number of the same one or several species in area *A-a* is the total individual number *N* of the

85   same one or several species in the total area *A*.

86

87   The fifth principle of sampling, species-area theory, is that the sum of the mathematical

88   expectation of number of species in area *a* and that of number of species lost if area *A-a* is

4

89   cleared is the total species number *M* in the total area *A*.

90

91   **Mathematical proof for the fourth and fifth principles of biodiversity sampling**

92   While the first, second and third principles of biodiversity sampling are straightforward and

93   easy to understand, a complementary method in a specific area is used to verify the fourth and

94   fifth principles.

95

96   The individual-area relationship (IAR) I_a in area *a* and $I_{A\text{-}a}$ in area *A-a*, for one sampling, if the

97   sampling area is *a* and the number of individuals in this area is *n*, there is always one-one

98   corresponding area of *N-n* individuals of the species in area *A-a* (*A* is the total area, and *N* is the

99   total individuals for one or several species in area *A*).

100   The mathematical expectation of individual-area relationship in area *a* is

$$ME(I_a) = \sum_{i=1}^{k} I_{a,i} P_i$$

101   where $ME(I_a)$ are the expected individuals for the individuals-area relationship in area *a* in

102   the *i*th sampling, *k* is the total samples, and the P_i is the possibility of the corresponding *i*th

103   sampling.

104   The mathematical expectation of individual-area relationship in area *A-a* is

$$ME(I_{A-a}) = \sum_{i=1}^{k} I_{A-a,i} Q_i$$

105   where $ME(I_{A-a})$ are the expected individuals for the individual-area relationship in area *A-a*

106   in the *i*th sampling, and the Q_i is the possibility of the corresponding *i*th sampling.

107   For any specific sampling *i*, there is always a one-one corresponding $I_{a,i} + I_{A-a,i} = N$, and the

5

108 $P_i = Q_i$ in this situation.

$$ME(I_a) + ME(I_{A-a}) = \sum_{i=1}^{k} I_{a,i}P_i + \sum_{i=1}^{k} I_{A-a,i}Q_i = \sum_{i=1}^{k}(I_{a,i} + I_{A-a,i}) * P_i = \sum_{i=1}^{k} NP_i = N\sum_{i=1}^{k} P_i$$

$$= N$$

109

110 The number of species in area a is m, then $S_a = m$, where species-area relationship (SAR) S_a is

111 the function of the species number with area a. The number of species lost in area A-a if the

112 area A-a is cleared is M-m, endemics-area relationship (EAR) E_{A-a} is the number of species

113 disappearing if area A-a is cleared.

114 The mathematical expectation of Species-area relationship in area a is

$$ME(S_a) = \sum_{i=1}^{k} S_{a,i}P_i$$

115 where $S_{a,i}$ is the species number in area a in the ith sampling, k is the total samples, and the P_i

116 is the possibility of the corresponding ith sampling.

117 The mathematical expectation of Endemics-area relationship in area A-a is

$$ME(E_{A-a}) = \sum_{i=1}^{k} E_{A-a,i}Q_i$$

118 where $E_{A-a,i}$ is the number of species disappear if area A-a is lost in the ith sampling, and the Q_i

119 is the possibility of the corresponding ith sampling.

120 For any specific sampling i, there is always a one-one corresponding $S_{a,i} + E_{A-a,i} = M$, and

121 the $P_i = Q_i$ in this situation.

$$ME(S_a) + ME(E_{A-a}) = \sum_{i=1}^{k} S_{a,i}P_i + \sum_{i=1}^{k} E_{A-a,i}Q_i = \sum_{i=1}^{k}(S_{a,i} + E_{A-a,i})P_i$$

$$= \sum_{i=1}^{k} MP_i = M\sum_{i=1}^{k} P_i = M$$

6

122

Conclusions

123

124 Sampling is a fundamental tool in ecology and critical for biodiversity measurement. However,

125 basic principles of biodiversity sampling have been overlooked for many years. In this paper, I

126 proposed and explored five principles of sampling for a specific area and biodiversity study.

127 The first principle of sampling, *species increasing with area*, is that the number of species

128 increases with the area. The second principle of sampling, *individuals increasing with area*, is

129 that the number of individuals increases with the area. The third principle of sampling, *sum of*

130 *species number*, is that the sum of species number in one area and species number in another

131 area is no less than the total species number in the two areas. The fourth principle of sampling,

132 *individual complement*, is that the sum of the mathematical expectation of individual number

133 of one or several species in the area *a* and that of the same one or several species in the area

134 *A-a* is the total individual number *N* of the same one or several species in the total area *A*. The

135 fifth principle of sampling, *species-area theory*, is that the sum of the mathematical expectation

136 of number of species in the area *a* and that of number of species lost if area *A-a* is cleared is the

137 total species number *M* in the total area *A*.

138

139 **Acknowledgments**

142

143 **References**

144   Achard F et al. 2002. Determination of deforestation rates of the world's humid tropical forests.

145   Science 297: 999–1002.

146   Connor EF, Courtney AC and Yoder JM. 2000. Individuals–area relationships: the relationship

147   between animal population density and area. Ecology 81: 734–748.

148   Green JL and Ostling A. 2003. Endemics-area relationships: the influence of species

149   dominance and spatial aggregation. Ecology 84: 3090-3097.

150   Green JL, Harte J and Ostling A. 2003. Species richness, endemism, and abundance patterns:

151   tests of two fractal models in a serpentine grassland. Ecology Letters 6: 919-928.

152   Harte J and Kinzig AP. 1997. On the implications of the species-area relationships for

153   endemism, spatial turnover, and food web patterns. Okios 80: 417-427.

154   Harte J, Kinzig AP and Green J. 1999. Self-similarity in the distribution and abundance of

155   species. Science 284: 334-336.

156   He F and Hubbell SP. 2011. Species-area relationships always overestimate extinction rates

157   from habitat loss. Nature 473: 368-371.

158   Heywood VH, Mace GM, May RM and Stuart SN. 1994. Uncertainties in extinction rates.

159   Nature 368: 105.

160   Kinzig AP and Harte J. 2000. Implications of endemics-area relationships for estimates of

161   species extinctions. Ecology 81: 3305-3311.

162   Mace GM, Gittleman JL and Purvis A. 2003. Preserving the tree of life. Science 300:

163   1707-1709.

164   Pan, XB. 2013. Fundamental equations for species-area theory. Scientific Reports. 2013, 3:
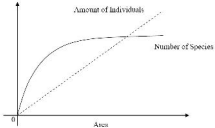
165   1334.

166 Pimm SL and Askins RA. 1995. Forest losses predict bird extinctions in eastern North America.

167 Proceedings of National Academic of Sciences 92: 9343–9347.

168 Rosenzweig ML. 1995. Species Diversity in Space and Time. Cambridge University Press,

169 Cambridge.

170 Thomas CD. et al. 2004. Extinction risk from climate change. Nature 427: 145-148.

171 Tilman D, May RM, Lehman CL and Nowak MA. 1994. Habitat destruction and the extinction

172 debt. Nature 371: 365–366.

173 Turner WR and Tjørve E. 2005. Scale-dependence in species-area relationships. Ecography 28:

174 721-730.

175 WCMC. 1992. Global Biodiversity: Status of the Earth's Living Resources. Chapman & Hall,

176 London.

177 World Resources Institute. 2005. Millennium Ecosystem Assessment. Ecosystems and Human

178 Well-Being: Biodiversity Synthesis.

179

180

181 **Figures Captions**

182 Figure 1. Relationship between the number of species and amount of individuals and the area

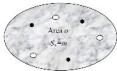183 Figure 2. Individual-area relationship I_a in area a and $I_{A\text{-}a}$ in area A-a.

184 Figure 3. Species-area relationship S_a in area a and endemics-area relationship $E_{A\text{-}a}$ in area A-a.

Amount of Individuals

Number of Species

0

Area

Area a

Individual number n in area a

Area $A-a$

Individual number $N-n$ in area $A-a$

Área x
$S = x_0$

Área $A-a$
$E_{A-a} = M-m$